

## RESEARCH ARTICLE

# Spatiotemporal video-domain high-fidelity simulation and realistic visualization of full-field dynamic responses of structures by a combination of high-spatial-resolution modal model and video motion manipulations

Yongchao Yang<sup>1</sup>  | Charles Dorn<sup>2</sup> | Tyler Mancini<sup>3</sup> | Zachary Talken<sup>4</sup> | Garrett Kenyon<sup>5</sup> | Charles Farrar<sup>1</sup> | David Mascareñas<sup>1</sup>

<sup>1</sup> Los Alamos National Lab, Engineering Institute, PO Box 1663, MS, Los Alamos, NM 87545, USA

<sup>2</sup> Department of Aerospace Engineering, California Institute of Technology, Pasadena, CA 91125, USA

<sup>3</sup> Department of Aerospace Engineering and Engineering Mechanics, University of Texas, Austin, TX 78712, USA

<sup>4</sup> Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA

<sup>5</sup> Applied Modern Physics, Los Alamos National Lab, PO Box 1663, MS D410, Los Alamos, NM 87545, USA

## Correspondence

Yongchao Yang, Argonne National Laboratory, 9700 S. Cass Ave., Lemont, IL 60439, USA.

Email: ycyang@anl.gov;  
yangyongchaohit@gmail.com

## Funding information

Los Alamos National Laboratory Lab Directed Research and Development (LDRD) program, Grant/Award Number: 20150708PRD2

## Summary

Structures with complex geometries, material properties, and boundary conditions exhibit spatially local dynamic behaviors. A high-spatial-resolution model of the structure is thus required for high-fidelity analysis, assessment, and prediction of the dynamic phenomena of the structure. The traditional approach is to build a highly refined finite element computer model for simulating and analyzing the structural dynamic phenomena based on detailed knowledge and explicit modeling of the structural physics such as geometries, materials properties, and boundary conditions. These physics information of the structure may not be available or accurately modeled in many cases, however. In addition, the simulation on the high-spatial-resolution structural model, with a massive number of degrees of freedom and system parameters, is computationally demanding. This study, on a proof-of-principle basis, proposes a novel alternative approach for spatiotemporal video-domain high-fidelity simulation and realistic visualization of full-field structural dynamics by an innovative combination of the fundamentals of structural dynamic modeling and the advanced video motion manipulation techniques. Specifically, a low-modal-dimensional yet high-spatial (pixel)-resolution (as many spatial points as the pixel number on the structure in the video frame) modal model is established in the spatiotemporal video domain with full-field modal parameters first estimated from line-of-sight video measurements of the operating structure. Then in order to simulate new dynamic response of the structure subject to a new force, the force is projected onto each modal domain, and the modal response is computed by solving each individual single-degree-of-freedom system in the modal domain. The simulated modal responses are then synthesized by the full-field mode shapes using modal superposition to obtain the simulated full-field structural dynamic response. Finally, the simulated structural dynamic response is embedded into the original video, replacing the original motion of the video, thus generating a new photo-realistic, physically accurate video that enables a

realistic, high-fidelity visualization/animation of the simulated full-field vibration of the structure. Laboratory experiments are conducted to validate the proposed method, and the error sources and limitations in practical implementations are also discussed. Compared with high-fidelity finite element computer model simulations of structural dynamics, the video-based simulation method removes the need to explicitly model the structure's physics. In addition, the photo-realistic, physically accurate simulated video provides a realistic visualization/animation of the full-field structural dynamic response, which was not traditionally available. These features of the proposed method should enable a new alternative to the traditional computer-aided finite element model simulation for high-fidelity simulating and realistically visualizing full-field structural dynamics in a relatively efficient and user-friendly manner.

**KEYWORDS**

full-field vibration measurement, high-fidelity simulation, modal analysis, motion synthesis, structural dynamics, video processing, visualization

## 1 | INTRODUCTION

Real-world operating structures are subjected to various dynamic loads. For example, civil infrastructure such as bridges and buildings confront operational loads (traffic and wind), natural hazards (earthquakes and hurricanes), and man-made extreme events (impacts and blasts). Aircrafts are subject to various aerodynamic and thermodynamic loads. Due to complex geometries, material properties, and boundary conditions, structures subjected to these loads exhibit spatially local dynamic behaviors. Therefore, a detailed knowledge of the structure (e.g., a high-spatial-resolution model) is required for high-fidelity analysis, assessment, and prediction of the dynamic phenomena of the structure.

The most common approach to achieve this is to build a detailed, highly refined, physics-based finite element computer model for simulating and analyzing the dynamic response of the structure based on detailed knowledge and explicit modeling of the structure's material properties, geometries, boundary conditions, joints, and so on.<sup>[1]</sup> Then the established finite element model (FEM) is usually updated by comparing the model-predicted response with the experimentally or field measured structural response.<sup>[2]</sup> Finally, one visualizes and animates the simulated dynamic behaviors on the computer model. While such an accurate, highly refined FEM of the structure is most desirable, it is usually extremely difficult to obtain.<sup>[3]</sup> For example, the properties of many civil structure materials such as concrete can seldom be accurately modeled. Such a challenge also exists in modeling the boundary conditions and component joints, especially when a detailed, refined model is required. In addition, updating and simulating the high-fidelity FEM, potentially with millions of degrees of freedom and system parameters, are computationally demanding to implement. Therefore, it is desirable to seek alternatives that enables high-fidelity modeling and simulating the dynamic behaviors of the structure in a more efficient and feasible manner.

A modal model, consisting of the natural frequencies, damping ratios, and mode shapes,<sup>[4]</sup> provides an effective approach to characterize the dynamics of the structure. Because structural dynamic responses can be modeled as a linear superposition of the modal responses, using mode shapes as expansion bases, structural dynamic responses at a high spatial resolution may be obtained if high-resolution mode shapes are available. An attractive feature of structural dynamics using a modal model is that structural dynamic responses are usually low dimensional in the modal domain,<sup>[4,5]</sup> that is, they can be approximated by a linear combination of a few dominant modes.

By using the modal transform, the original problem—simulating and computing the structural dynamic responses by solving the coupled multi-degree-of-freedom equations of the spatial model—can be decomposed into a much simpler problem of solving a set of a few individual single-degree-of-freedom (SDOF) systems. The latter approach is more computationally efficient. Therefore, a modal model with a few dominant modes and high-resolution mode shapes is highly feasible for high-fidelity modeling and simulating structural dynamics.

Building a high-spatial-resolution modal model for high-fidelity simulating dynamic responses of the structure requires knowing the natural frequencies, damping ratios, and high-resolution mode shapes, which can be identified

from vibration measurements by experimental or operational modal analysis methods. Experimental modal analysis methods use both measured inputs (excitation) and outputs (response) to the structure,<sup>[4]</sup> while operational or output-only modal analysis methods rely only on response measurements.<sup>[6]</sup> Operational modal analysis is considered a useful alternative when controllable excitation is difficult to apply on the structure or the operating excitation is challenging to measure, especially for larger scale structures.

Traditional vibration measurement techniques such as using accelerometers are inadequate for high-spatial-resolution modal analysis, because these discrete point-wise sensors only provide sparse, low-spatial sensing resolution measurements, while their dense deployment to achieve high-spatial resolution is not feasible due to high costs, mass-loading effect, and significant surface modification on the structure. Many noncontact measurement techniques, such as those using scanning laser vibrometers,<sup>[7]</sup> provide high spatial resolution displacement measurements without the need to install sensors on the structure or inducing the mass-loading effect. However, they are relatively expensive and make measurements sequentially, which requires intensive time and labor for measurements of larger area. As an alternative noncontact measurement method, digital video cameras are relatively low cost, agile, and provide simultaneous, high-spatial resolution measurements. Combined with image processing algorithms (e.g., image correlation<sup>[8]</sup> or point tracking and optical flow<sup>[9]</sup>), video camera-based measurements have been successfully used for structural dynamic response measurement<sup>[10–15]</sup> and experimental modal analysis<sup>[16–21]</sup> to obtain full-field mode shapes. One issue associated with these methods is that they typically require speckle pattern or high-contrast markers to be placed on the surface of structures, which induces the surface preparation and target installation issue and is less feasible when the measurement area is large or inaccessible. Recently, the phase-based video motion estimation technique<sup>[22–24]</sup> has been explored to perform output-only modal identification without the need to install high-contrast markers or speckle paints on the structure's surface.<sup>[25,26]</sup> The lately developed algorithm based on this technique enables extraction and visualization of full-field modal parameters in a relatively efficient and automated manner,<sup>[26]</sup> and is suited for establishing a full-field modal model and further for realistic visualization of the simulated full-field structural dynamic response in this study.

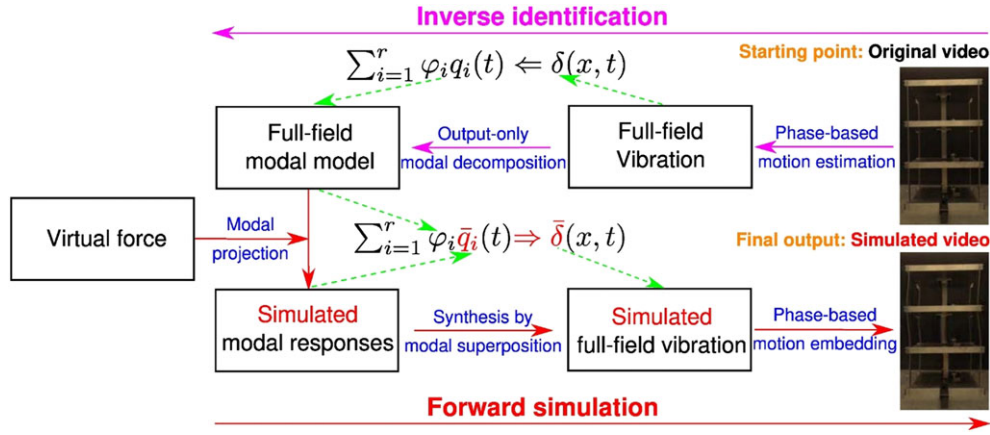
Aiming for an alternative to the finite element computer model simulation, this study establishes a new video-based method for high-fidelity simulation and realistic visualization of full-field structural dynamic response using a combination of the full-field modal model of the structure and video motion manipulation techniques. First, the full-field modal parameters are estimated from line of sight video measurements by the previously developed video phase-based output-only modal identification method, and the full-field modal model is established in the video domain. Then, in order to simulate new dynamic response of the structure subject to new force, the force is projected onto each modal domain, and the modal response is computed by solving each individual SDOF system in the modal domain. The simulated modal responses are then synthesized by the full-field mode shapes using modal superposition to obtain the simulated full-field structural responses. Finally, the simulated full-field structural responses are embedded into the original video, replacing its original motion, to generate a new, photo-realistic, physically accurate video, which enables full-field, high-fidelity, and realistic visualization/animation of the simulated motion in a realistic video of the structure. The last step is relevant to the motion synthesis technique in computer graphics and animation.<sup>[27–29]</sup>

Compared with finite element computer model simulations that require detailed information of the structure's physics, the video-based simulations implicitly incorporates these information in the modal model. In addition, the low-modal-dimensional and decoupled nature of the modal model allows the proposed approach to directly solve only a few SDOF system equations in the modal domain to obtain the structural dynamic response, which is computationally efficient. Finally, the simulated full-field structural dynamic responses can be realistically visualized and animated in a photo-realistic, physically accurate video of the structure, as compared with those cartoon-like figures created by the finite element computer model simulations.

The remainder of the paper is organized as follows. Section 2 introduces the theoretical background of the proposed video-based method for high-fidelity simulation and realistic visualization of full-field structural dynamics based on a combination of the established full-field modal model and video motion manipulation techniques. Section 3 validates the proposed method on a proof-of-principle basis by laboratory experiments on a three-story building structure. Section 4 presents conclusions and future work.

## 2 | APPROACH

The new approach for high-fidelity simulation and realistic visualization of structural dynamics is based on a novel combination of the full-field modal model of the structure and video motion manipulation techniques. Figure 1 shows



**FIGURE 1** The flowchart of the proposed video-based approach for high-fidelity simulation and realistic visualization of full-field structural dynamic response. Step 1 “inverse identification” with pink arrows: establishing a full-field modal model estimated from original video measurements (top right). It first estimates the full-field vibration (Equation (1)–(2)) and then the full-field modal parameters (Equation (3)–(13)). Step 2 “forward simulation” with red arrows: synthesizing and embedding the simulated full-field structural responses into the original video to generate a new video, enabling realistic visualization of full-field dynamic response. It first projects the equation of motion with the virtual force onto the modal domain (Equation (14)–(20)) and computes the simulated modal responses by solving Equation (21). Then the simulated modal responses are synthesized by the full-field mode shapes using Equation (25) to produce the simulated full-field structural responses, which are finally embedded into the original video to generate a new video (Equation (26))

a flowchart of the new approach, which mainly consists of two steps. The first step “inverse identification” is to first establish a full-field modal model of the structure in the video domain from the original video measurements of the structure, which is introduced in the following Section 2.1. The second step “forward simulation” detailed in Section 2.2 introduces the procedures of simulating and visualizing the new full-field dynamic response in a newly-generated photo-realistic, physically accurate video of the structure.

## 2.1 | Establishment of the full-field modal model in video domain

A full-field modal model is first established consisting of modal frequencies, damping ratios, and full-field mode shapes, which are estimated from line of sight video measurements of the operating structure. Note that this part of the proposed method has been detailed in the authors' recent work<sup>[26]</sup>; it is reviewed here from Section 2.1.1 to Section 2.1.3.

### 2.1.1 | High-spatial-resolution vibration motion in video measurement

Structural vibration can be measured into video records containing frames with temporally displaced or translated image intensity  $I(x + \delta(x, t))$  (assuming  $N$  pixels and  $T$  frames), where  $x$  is the spatial variable (pixel coordinate),  $t$  is the temporal variable, and  $\delta(x, t)$  is the spatially local, temporally varying motion. Just as oscillating motion can be characterized by its magnitude and phase using Fourier representation, structural vibration motion is encoded in the local amplitudes  $\rho(x, t)$  and local phases  $\psi(x, t)$  of  $I(x + \delta(x, t))$ , which can be extracted from  $I(x + \delta(x, t))$  by the spatially multiscale, localized, and complex steerable pyramid filters<sup>[23,30]</sup> (complex Gabor filters, which are sinusoids windowed by a Gaussian envelop leading to finite spatial support):

$$I(x + \delta(x, t)) = \sum_{\omega=-\infty}^{\infty} R_{\omega}(x, t) = \sum_{\omega=-\infty}^{\infty} \rho_{\omega}(x, t) e^{j2\pi\omega(x+\delta(x, t))}, \quad (1)$$

in which  $R_{\omega}(x, t)$  is the subband representation (filter response) at the spatial scale  $\omega$

$$R_{\omega}(x, t) = \rho_{\omega}(x, t) e^{j2\pi\omega(x+\delta(x, t))}, \quad (2)$$

then local phases  $\psi_{\omega}(x, t) = 2\pi\omega x + 2\pi\omega\delta(x, t)$  and local amplitude  $\rho_{\omega}(x, t)$  are obtained from  $R_{\omega}(x, t)$ . Therefore, the normalized  $\psi_{\omega}(x, t)$  directly correspond to the vibration motion  $\delta(x, t)$  and have been shown to be relatively insensitive

to changes in illumination, perspective, and surface conditions,<sup>[22]</sup> making it preferable over image intensity information in many real-world applications.

The obtained local phase  $\psi(x, t)$  corresponding to the vibration motion  $\delta(x, t)$  can be expressed by the modal expansion as linear combinations of the modal responses

$$\delta(x, t) = \Phi(x)q(t) = \sum_{i=1}^n \varphi_i(x)q_i(t), \quad (3)$$

where  $\Phi \in \mathbb{R}^{N \times n}$  and  $q \in \mathbb{R}^{n \times T}$  with  $\varphi_i(x)$  as the  $i$ th mode shape and  $q_i(t)$  as the  $i$ th modal coordinate, both of which need to be identified from the only knowledge of the vibration  $\delta(x, t)$ . Because the pixel dimension of  $\delta(x, t)$ ,  $N$ , is much higher than the mode dimension (number),  $n$ , that is,  $N \gg n$ , the output-only modal identification problem above is overcomplete. Therefore, spatial dimension reduction of the overcomplete system described by Equation (3) is first conducted by performing principal component analysis (PCA) on  $\delta \in \mathbb{R}^{N \times T}$ , as detailed in the following section.

### 2.1.2 | Low-modal-dimensional structural dynamics

PCA is closely related to the singular value decomposition of the motion matrix  $\delta \in \mathbb{R}^{N \times T}$  (assuming  $N > T$ )

$$\delta = U\Sigma V^* = \sum_{i=1}^n \sigma_i u_i v_i^*, \quad (4)$$

where  $\Sigma \in \mathbb{R}^{N \times T}$  is a diagonal matrix containing  $T$  nondecreasing diagonal elements,  $\sigma_i$  as the  $i$ th singular value ( $\sigma_1 \geq \dots \geq \sigma_i \geq \dots \geq \sigma_T \geq 0$ ), and  $U = [u_1, \dots, u_N] \in \mathbb{R}^{N \times N}$  and  $V = [v_1, \dots, v_T] \in \mathbb{R}^{T \times T}$  (“\*” denotes “transpose”) are the matrices of the left-singular and right-singular vectors obtained by the eigenvalue decomposition of the covariance matrices of  $\delta$

$$\delta\delta^* = U\Sigma^2 U^*, \quad (5)$$

$$\delta^*\delta = V\Sigma^2 V^*. \quad (6)$$

The rank of  $\delta$  is  $r$  if the number of nonzero singular values is  $r$  ( $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_T = 0$ ).

It is known that the  $i$ th singular value  $\sigma_i$  is related to the energy projected onto the  $i$ th principal direction (vector)  $u_i$  of  $\delta$ . In structural dynamics, for an undamped or very lightly damped structure, it was shown that if its mass matrix is proportional to the identity matrix, the principal directions  $u_i$  ( $i = 1, \dots, N$ ) will converge to the mode shape directions<sup>[31]</sup> with the corresponding singular values  $\sigma_i$  ( $i = 1, \dots, N$ ) indicating their participating energy in the structural vibration responses  $\delta$ . In other words, the structure's active modes in  $\delta$ , under broadband excitation, are projected onto the  $r \approx n$  principal components. An empirical but usually true observation is that typically there are only few dominant active modes present in the structural vibration responses. This means that the rank of  $\delta$ ,  $r$ , which is approximately the number of active singular values ( $\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} \approx \dots \approx \sigma_T \approx 0$ ), is very small compared with the spatial dimension of  $\delta$ ,  $N$ ; that is,  $r \ll N$ . Therefore, PCA is able to conduct significant spatial dimension reduction on  $\delta$  by linearly projecting most energy (modal components) of  $\delta$  onto a very small number of principal components

$$\eta = U_r^* \delta, \quad (7)$$

where  $U_r = [u_1, \dots, u_r] \in \mathbb{R}^{N \times r}$  is the first  $r$  ( $\ll N$ ) columns (orthonormal) of  $U$ , and the  $i$ th row of the resultant  $\eta = [\eta_1, \dots, \eta_r]^* \in \mathbb{R}^{r \times T}$ ,  $\eta_i$ , is the  $i$ th principal component of  $\delta$ . Since the low-rank  $\delta \in \mathbb{R}^{N \times T}$  can be recovered by linearly projecting  $\eta \in \mathbb{R}^{r \times T}$  back to

$$\delta \approx U_r \eta. \quad (8)$$

### 2.1.3 | Blind mode separation of principal components

Performing PCA on  $\delta \in \mathbb{R}^{N \times T}$  projects the active mode components within  $\delta \in \mathbb{R}^{N \times T}$  onto  $r \approx n$  principal components  $\eta \in \mathbb{R}^{r \times T}$ , which can also be expressed as linear mixtures of the modal coordinates



$$\eta(x, t) = Y(x)q(t) = \sum_{i=1}^r \gamma_i(x)q_i(t), \quad (9)$$

where  $Y(x)$  is the mixing matrix in the “principal component” domain. Substituting Equation (3) and Equation (9) into Equation (8) yields

$$\sum_{i=1}^n \varphi_i q_i = \delta \approx U_r \eta = U_r Y q = U_r \sum_{i=1}^r \gamma_i q_i = \sum_{i=1}^r (U_r \gamma_i) q_i. \quad (10)$$

Comparing the two ends of Equation (10) with  $r \approx n$ , it yields

$$\varphi_i \approx U_r \gamma_i, \quad i = 1, \dots, r. \quad (11)$$

It has been recently established<sup>[32–41]</sup> that there is a one-to-one mapping between the modal superposition model and the linear mixture model of the blind source separation (BSS), which can perform output-only modal identification efficiently. Among a family of BSS techniques that are suited for output-only modal identification, the complexity pursuit (CP) algorithm<sup>[36]</sup> was shown to be efficient and able to identify closely-spaced and highly-damped modes with little expert supervision and parameter adjustments; it is therefore adopted in this study. Therefore, using CP to directly decouple or separate  $\eta(x, t)$  into individual modal coordinates

$$q(t) = W(x)\eta(x, t), \quad (12)$$

simultaneously yielding the demixing or decoupled matrix  $W \in \mathbb{R}^{r \times r}$  and the modal coordinates  $q(t)$ . Comparing Equation (9) with Equation (12),  $Y \in \mathbb{R}^{r \times r}$  is the inverse of  $W$ .

$$Y = W^{-1} \quad (13)$$

Therefore, the high-resolution mode shapes  $\varphi_i(x)$  ( $i = 1, \dots, r$ ) are estimated as per Equation (11) to Equation (13). Modal frequencies  $\omega_i$  and damping ratios  $\zeta_i$  can be estimated from the obtained modal coordinates  $q_i(t)$  ( $i = 1, \dots, r$ ) by Fourier transform and logarithm decrement techniques, respectively, or using Hilbert transform (HT)-based SDOF analysis methods.<sup>[42]</sup> With the modal frequencies  $\omega_i$ , damping ratios  $\zeta_i$ , and full-field mode shapes  $\varphi_i(x)$  ( $i = 1, \dots, r$ ) estimated from the line of sight video measurements, the full-field modal model is established in the video domain where the full-field mode shapes (as many spatial points as the pixel number on the structure in the video frame) are in the pixel coordinate.

## 2.2 | High-fidelity simulation and realistic visualization of full-field dynamic response

The next step is to simulate new structural dynamic response subject to new force with the established full-field modal model. Using the low-modal-dimensional and decoupled nature of the modal model, the modal response is first computed in each of the few dominant modal domain by solving an SDOF system (detailed in Section 2.2.1) and then synthesized by the full-field mode shapes using modal superposition to produce the full-field structural dynamic response (Section 3). Because the mode shapes are estimated from output only, they are nonscaled; this issue will be addressed in the Section 2.2.2, assuming a partial knowledge of the mass distribution of the structure is available.

### 2.2.1 | Simulating modal response to the modally projected force

#### Modal projection

The spatial model of the equation of motion of a linear time-invariant system subject to a new simulated force  $\bar{f}(t)$  (the upper bar “—” denotes “simulated”) is

$$M\ddot{\bar{\delta}}(t) + K\bar{\delta}(t) + C\dot{\bar{\delta}}(t) = \bar{f}(t), \quad (14)$$

where  $M$ ,  $K$ , and  $C$  are the diagonal mass, symmetric stiffness, and symmetric damping matrices (assuming proportional damping), respectively. Using the modal expansion of the motion Equation (3) and projecting Equation (14) onto the modal domain

$$\Phi^* M \Phi \ddot{\bar{q}}(t) + \Phi^* K \Phi \bar{q}(t) + \Phi^* C \Phi \dot{\bar{q}}(t) = \Phi^* \bar{f}(t), \quad (15)$$

the motion is decoupled into  $n$  SDOF systems

$$m_i^* \ddot{\bar{q}}_i(t) + k_i^* \bar{q}_i(t) + c_i^* \dot{\bar{q}}_i(t) = \bar{f}_i^*(t), \quad i = 1, \dots, n, \quad (16)$$

where “ $\star$ ” denotes “modal” (not to confuse with the “transpose” “ $*$ ”), and the modal mass, modal stiffness, modal damping matrices, and modal force are, respectively,

$$M^* = \Phi^* M \Phi, \quad m_i^* = \varphi_i^* M \varphi_i, \quad i = 1, \dots, n, \quad (17)$$

$$K^* = \Phi^* K \Phi, \quad k_i^* = \varphi_i^* K \varphi_i, \quad i = 1, \dots, n, \quad (18)$$

$$C^* = \Phi^* C \Phi, \quad c_i^* = \varphi_i^* C \varphi_i, \quad i = 1, \dots, n, \quad (19)$$

$$\bar{f}^*(t) = \Phi^* \bar{f}(t), \quad \bar{f}_i^*(t) = \varphi_i^* \bar{f}(t), \quad i = 1, \dots, n. \quad (20)$$

Because  $k_i^* = m_i^* \omega_i^2$  and  $c_i^* = 2\zeta_i m_i^* \omega_i$  (the modal frequency  $\omega_i$ , damping ratio  $\zeta_i$ , and mode shape  $\varphi_i$  have been estimated from the video measurements), one only needs to estimate the modal mass  $m_i^*$  (to be addressed in Section 2.2.2) in Equation (16) to compute the modal response  $\bar{q}_i(t)$  to the projected modal force  $\bar{f}_i^*(t)$ .

$$m_i^* \ddot{\bar{q}}_i(t) + m_i^* \omega_i^2 \bar{q}_i(t) + 2\zeta_i m_i^* \omega_i \dot{\bar{q}}_i(t) = \varphi_i^* \bar{f}(t), \quad i = 1, \dots, n \quad (21)$$

Note that the modal dimension  $n$  is far smaller than the spatial dimension  $N$ ; therefore, based on the principle of modal transform (using the estimated modal parameters), one only needs to solve  $n \ll N$  decoupled SDOF system equations (Equation (21)), instead of the coupled  $N$ -degree of freedom (DOF) system of equations (Equation (14)).

### Estimation of the modal mass

Solving Equation (21) requires estimating the modal mass  $m_i^*$  ( $i = 1, \dots, n$ ). In theory, because the spatial dimension  $N$  is very high, the spatial model Equation (14) is highly refined; a large-scale mass matrix  $M$  ( $N$  diagonal mass elements) is therefore needed to estimate the modal mass  $m_i^*$  according to Equation (17). In the following, however, it will be derived that the highly refined spatial model with a massive amount of mass elements can be dramatically reduced for estimating  $m_i^*$ .

Expanding Equation (17),

$$m_i^* = \varphi_i^* M \varphi_i = \sum_{j=1}^N \varphi_{ji} m_j \varphi_{ji} = \sum_{j=1}^N m_j \varphi_{ji}^2, \quad (22)$$

where  $m_j$  is the  $j$ th diagonal element (mass DOF) of  $M$ , and  $\varphi_{ji}$  is the  $j$ th element of  $\varphi_i \in \mathbb{R}^N$ . For building structures, the motion of the floors, which take up most mass of the building, are rigid; for example, the mode shape component of all the mass elements at the  $k$ th floor ( $k = 1, \dots, K$ ) would approximately be the same as  $(\varphi_i)_k$ . Therefore, Equation (20) becomes

$$m_i^* = \sum_{j=1}^N m_j \varphi_{ji}^2 = \sum_{k=1}^K (m)_k (\varphi_i)_k^2 + \sum_{j \notin [1, K]} m_j \varphi_{ji}^2, \quad (23)$$

where  $(m)_k$  is the cumulative mass of all elements at the  $k$ th floor with a mode shape component  $(\varphi_i)_k$ . Because  $(m)_k$  ( $k = 1, \dots, K$ ) are dominant and the rest of the mass elements,  $m_j$  ( $j \notin [1, K]$ ) are very small, so Equation (23) can be

approximated by ignoring those associated with the small  $m_j$   $\left(j \neq [1 K]\right)$ .

$$m_i^* = \sum_{j=1}^N m_j \varphi_{ji}^2 \approx \sum_{k=1}^K (m)_k (\varphi_i)_k^2, \quad (24)$$

which substantially reduces the highly refined  $N$ -dimensional spatial model to a small  $K$ -dimensional model ( $K \ll N$ ) for estimation of the modal mass  $m_i^*$ . Note that the above approximation is true only when the structure can be approximately modeled as lumped-mass type with rigid motion; if not, then the knowledge of the full mass matrix is needed to estimate the modal masses.

### 2.2.2 | Synthesizing simulated modal responses to full-field dynamic response

After obtaining the simulated modal response  $\bar{q}_i(t)$  ( $i = 1, \dots, n$ ) to the projected modal force  $\bar{f}_i^*(t)$  by solving Equation (21), the simulated structural responses  $\bar{\delta}(x, t)$  to the force  $\bar{f}(t)$  can be obtained by synthesizing  $\bar{q}_i(t)$  ( $i = 1, \dots, n$ ) by the full-field mode shapes  $\varphi_i(x)$  ( $i = 1, \dots, r$ ) using modal superposition

$$\bar{\delta}(x, t) = \Phi(x) \bar{q}(t) = \sum_{i=1}^n \varphi_i(x) \bar{q}_i(t), \quad (25)$$

which projects the simulated modal response  $\bar{q}_i(t)$  in the modal domain back to the full-field simulated structural response  $\bar{\delta}(x, t)$  in spatio (pixel)-temporal domain.

### 2.2.3 | Realistic visualization of simulated full-field dynamic response in video

Finally, the simulated full-field structural dynamic response  $\bar{\delta}(x, t)$  is embedded into the original video  $I(x + \delta(x, t))$  but replacing the original motion  $\delta(x, t)$  with the new simulated motion  $\bar{\delta}(x, t)$ , finally generating a new video consisting of images  $\bar{I}(x + \bar{\delta}(x, t))$ :

$$\bar{I}(x + \bar{\delta}(x, t)) = \sum_{\omega=-\infty}^{\infty} \bar{R}_{\omega}(x, t) = \sum_{\omega=-\infty}^{\infty} \rho_{\omega}(x, t) e^{j2\pi\omega(x + \bar{\delta}(x, t))}, \quad (26)$$

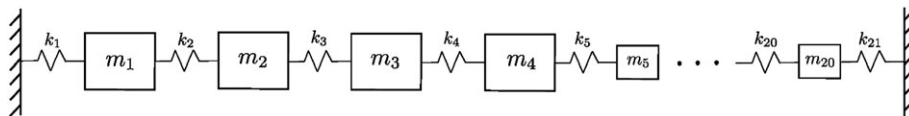
which is implemented by an *inverse* complex steerable filtering (mentioned in Section 2.1.1 and detailed in Yang et al.<sup>[26]</sup>) with replacement of  $\delta(x, t)$  by  $\bar{\delta}(x, t)$ . Therefore, this new simulated video  $\bar{I}(x + \bar{\delta}(x, t))$  enables photo-realistic, physically accurate visualization and animation of the simulated full-field structural dynamic responses  $\bar{\delta}(x, t)$  (at each involved pixel) in a realistic video of the structure.

In the following, a numerical model is used to validate the above procedures presented in Section 2.1 to Section 2.2.

## 2.3 | Validations on an overcomplete numerical model

A numerical study is presented in this section to demonstrate the procedures for establishing the modal model and synthesizing the simulated motion. It also serves to validate the proposed PCA-CP method for output-only modal identification and the modal mass estimation scheme in an overcomplete setting. A 20-DOF mass-spring-damper system (Figure 2) was constructed with the following parameters: diagonal mass matrix  $M$  with dominant masses  $m_1 = 200$ ,  $m_2 = m_3 = m_4 = 150$  and small masses  $m_5 = \dots = m_{20} = 1$ , stiffness  $k_1 = \dots = k_{21} = 5000$ , and proportional damping matrix  $C = \alpha M$  with  $\alpha = 0.03$ .

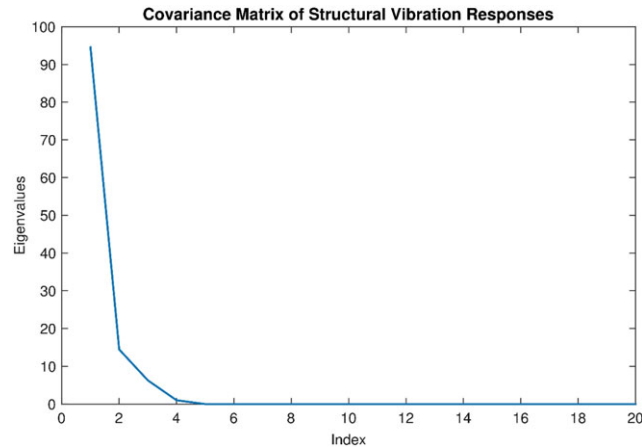
First, the modal model was established. When performing the output-only modal identification, an initial unit velocity at the first DOF was used to excite the structure and induced free vibration. Twenty-second 20-channel structural



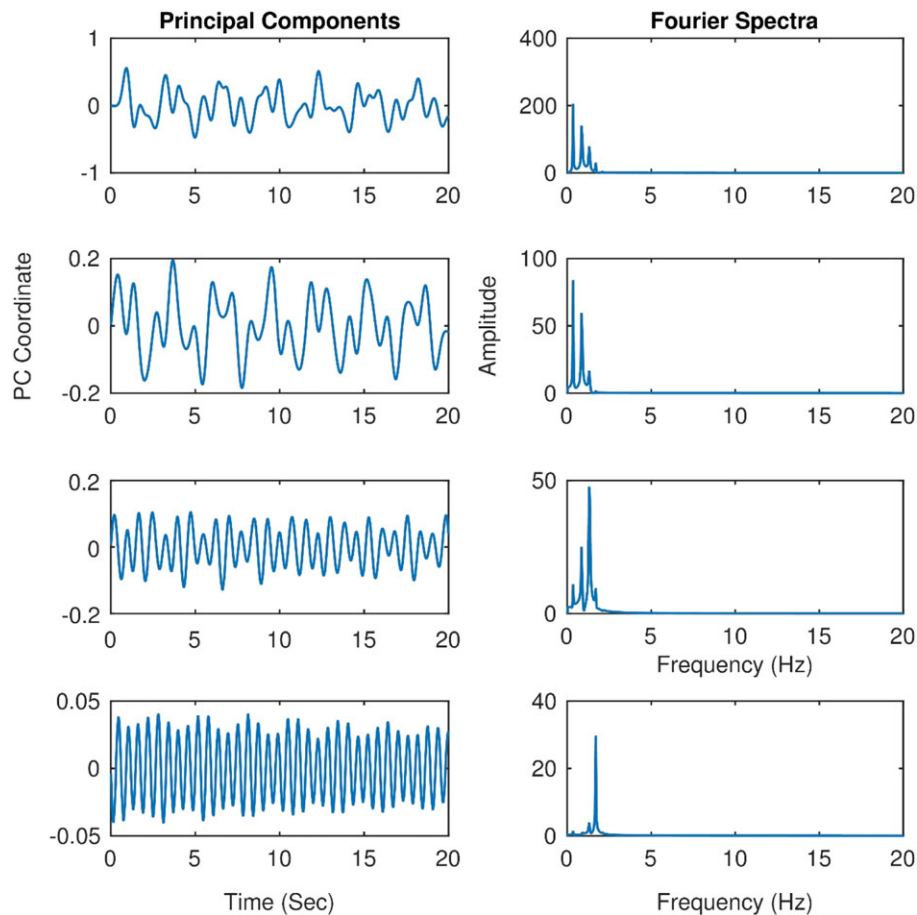
**FIGURE 2** A spring-mass damped numerical model with 20 degrees of freedom where four of them are very dominant



responses are computed by the Newmark-beta algorithm with a sampling frequency of 100 Hz. The PCA-CP algorithm was then applied on the structural responses as per Equation (3) to Equation (13). Figure 3 shows that there are only four active nonzero eigenvalues (squares of the singular values) in the covariance matrix of the structural responses, indicating most of the modal components contained in the 20-channel have been projected on to the first four principal components, which is depicted in Figure 4. It has also been shown in Figure 4 that while each of the four principal components contains 1 dominant mode of the four modes, respectively, it is still a mixture of the four modes. After applying CP on these four principal components according to Equation (12), they are completely separated into four individual



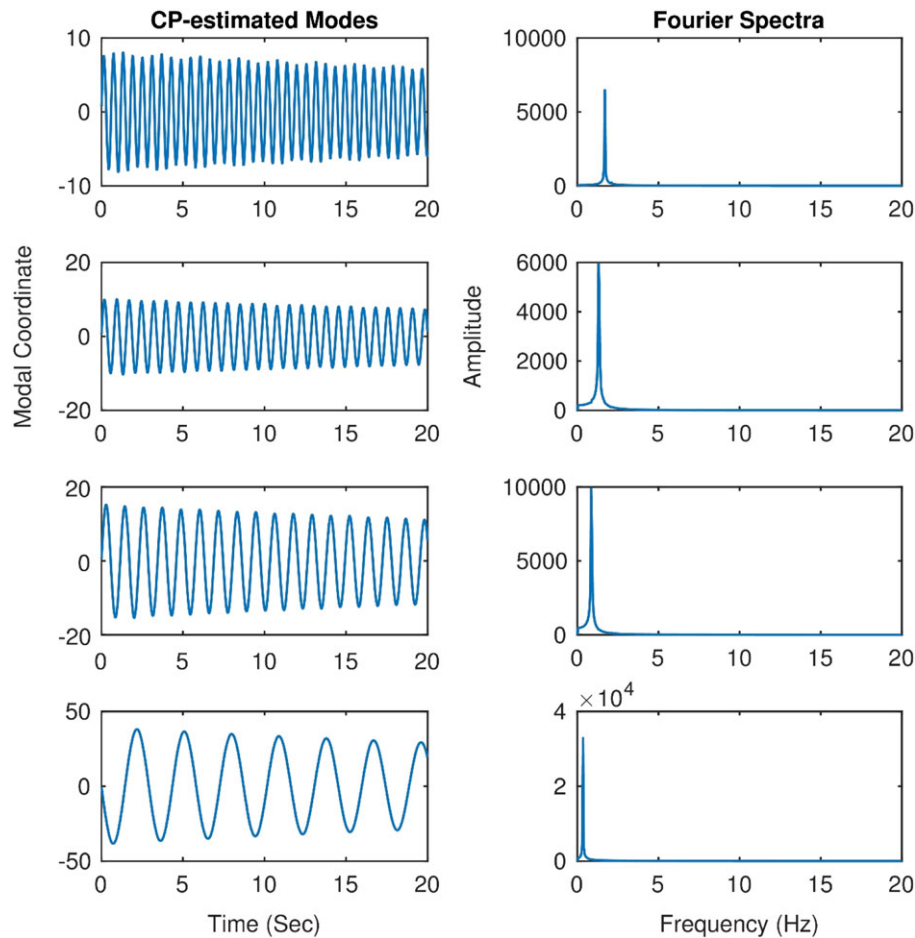
**FIGURE 3** The eigenvalue distribution of the covariance matrix of the structural responses of the 20-degree of freedom numerical model



**FIGURE 4** The principal components and their Fourier spectra of the structural responses of the 20-degree of freedom numerical model

modal coordinates (shown in Figure 5), upon which HT-based SDOF analysis methods are performed to extract the modal frequencies and damping ratios. The mode shapes were estimated according to Equation (11) and correlated with the analytical ones using the modal assurance criterion. Table 1 lists the estimated modal parameters that match the analytical ones very well. The modal masses of these four modes were then estimated using Equation (24) with the estimated mode shapes and the knowledge of the four dominant masses (ignoring all those small masses whose knowledge usually not available). The estimated modal masses are seen very close to the analytical ones. Thus, a modal model with the estimated modal parameters was established.

Next, using the established modal model, simulated structural responses to the Chichi earthquake excitation<sup>[43]</sup> were generated as follows. For each of the four dominant modes, the earthquake excitation applied at the first DOF was



**FIGURE 5** The modal coordinates estimated by the complexity pursuit (CP) algorithm applied on the active principal components of the structural responses of the 20-degree of freedom numerical model

**TABLE 1** Modal parameters of an overcomplete numerical system estimated by PCA-CP

Mode	Modal frequency (Hz)		Damping ratio (%)		MAC (mode shape)	Modal mass	
	Theoretical	Estimated	Theoretical	Estimated		Theoretical	Estimated
1	0.3444	0.3442	0.69	0.69	1.0000	383.38	375.24
2	0.8709	0.8706	0.27	0.29	1.0000	434.33	424.00
3	1.3239	1.3234	0.18	0.19	1.0000	535.29	533.24
4	1.6919	1.6911	0.14	0.13	0.9997	356.06	344.41

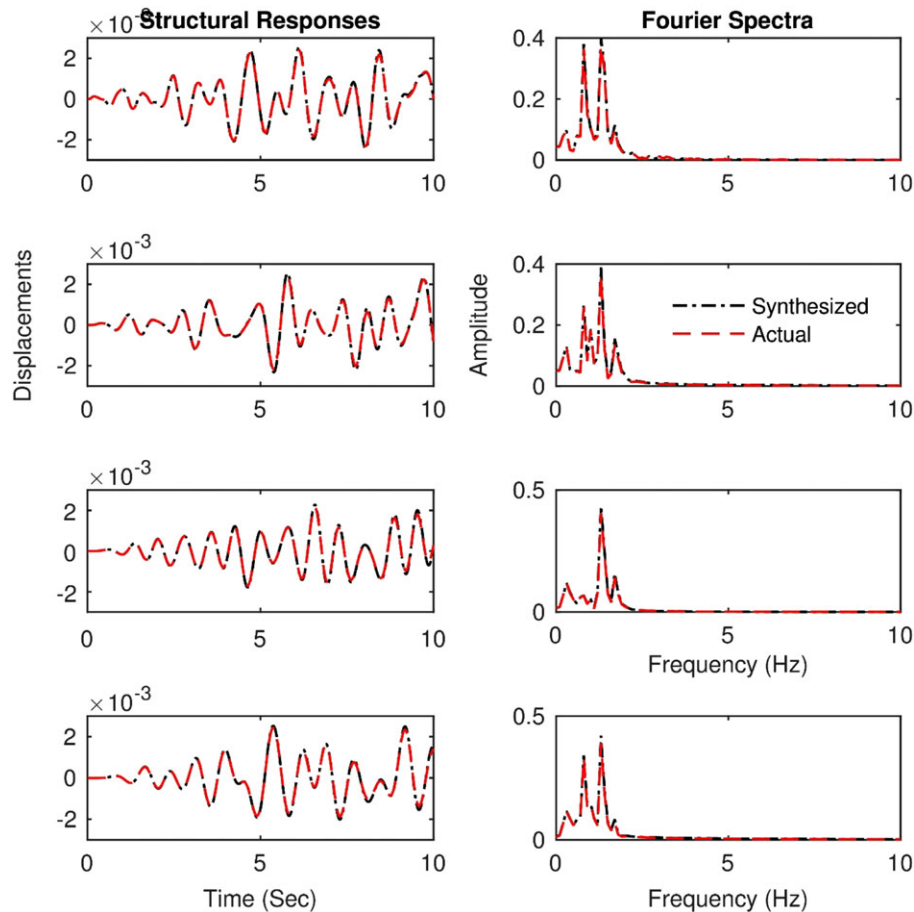
Note: MAC = modal assurance criterion; PCA-CP = principal component analysis-complexity pursuit.

projected onto the modal domain using the corresponding estimated mode shape. Then the modal responses are computed by solving the SDOF models Equation (21) using the Newmark-beta algorithm. Then these modal responses were synthesized by the mode shapes using the modal superposition Equation (25) to obtain the structural responses. These simulated structural responses were compared with the actual earthquake responses generated from the full 20-DOF system. Figure 6 shows that they correlate with each other quite well, indicating that using the estimated modal model consisting of few dominant modes and the knowledge of few dominant rigid masses are able to predict the structural responses of the full (spatial) model well.

## 2.4 | Summary of the video-based simulation method

The proposed video-based method for high-fidelity simulation and realistic visualization of full-field structural dynamics is capable of embedding the arbitrary forces induced full-field structural dynamic responses predicted by an established full-field modal model into a new video of the structure by video motion synthesis and manipulations. Summarizing those presented as per Equation (1) to Equation (26), the proposed method whose flowchart shown in Figure 1 undergoes the following steps:

Step 1. “inverse identification”: (a) Establish the full-field modal model of the structure by estimating the full-field modal parameters using the motion extracted from line of sight video measurement of the operating (output only) structure as per Equation (1) to Equation (13) and computing the modal masses as per Equation (22) to Equation (24).



**FIGURE 6** The synthesized structural responses and their Fourier spectra using a modal model consisting of only the first four modes, compared with those directly produced by the full 20-degree of freedom numerical model. (showing only the first four channels' displacements. Black point dash line: synthesized; red dash line: actual)

Step 2. “forward simulation”: (b) Project the force onto the modal domain and compute the corresponding modal responses by solving Equation (21). Obtain the full-field structural responses by synthesizing the modal responses using modal superposition Equation (25).

(c) Obtain a new photo-realistic, physically accurate video embedded with the full-field synthesized dynamic response (at each involved pixel) by replacing the original video motion with the synthesized dynamic response and performing the inverse transform (reconstruction) of the complex steerable filters (Equation (26)).

## 2.5 | Comparisons with computer-aided FEM simulation and visualization

Compared with finite element computer model simulations that require detailed specification of the structure's geometries, materials properties, and boundary conditions, the video-domain modal model implicitly incorporates these information in its constituent full-field modal parameters that are estimated from line of sight video measurements of the operating structure. In addition, the low-modal-dimensional and decoupled nature of the full-field modal model allows the proposed approach to directly solve only few SDOF system equations in the modal domain to obtain the structural dynamic response. In contrast, traditional finite element computer model simulations require either solving the high-spatial-density, coupled multi-degree-of-freedom system of equations directly or first finding the eigensolutions for the large-scale spatial model to solve for the dynamic structural responses in the modal domain. The final results of the proposed procedures can be visualized and animated in a photo-realistic, physically accurate video of the structure, as compared with those cartoon-like figures in a computer model of the finite element computer model simulation. Table 2 lists a detailed comparisons of the two approaches for high-fidelity simulation and visualization of structural dynamic responses.

## 3 | EXPERIMENTAL VALIDATION

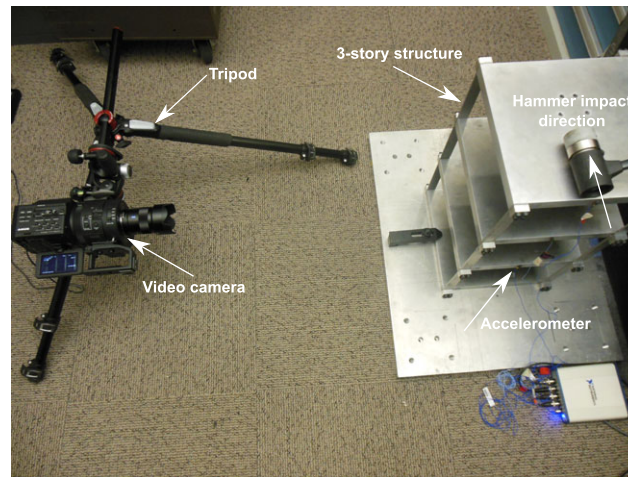
### 3.1 | Experimental setup

In this section, the proposed method is validated by laboratory experiments conducted on a bench-scale model of a three-story building structure. The structure consists of aluminum columns and lumped mass plates on each floor, with its base fixed. A stationary camera (Sony NXCAM with a pixel resolution of  $1920 \times 1080$ ) outfitted with a Zeiss lens with a fixed focal length of 24 mm was used to take line of sight video measurements of the structure. The frame rate was set at 240 frames per second. The illumination environment is the ordinary indoor lighting condition without any external

**TABLE 2** Comparisons of the features of the two approaches for high-fidelity simulation and visualization of structural dynamics

Comparisons	Finite element computer model simulation	Video-domain modal model simulation
Model	Spatial model requiring detailed specifications of the structure's physics (geometries, boundary conditions, material properties, etc).	Modal model with full-field modal parameters estimated from line of sight video measurements of the structure
Dimension	Usually high-spatial dimension	Usually low modal dimension
Computation	Solve large-scale coupled multi-DOF system equations —or— Perform eigen-analysis on the large-scale spatial model for reduced-order modal domain single-DOF system computation	Directly solve very few decoupled single-DOF system equations
Results	Cartoon-like visualization and animation of structural dynamic behaviors in the computer model	Photo-realistic visualization and animation of structural dynamic behaviors in a realistic video of the structure
Errors	In modeling the structure's physics, solving the large-scale coupled system equations, and so on	In estimating/establishing the full-field modal model, representing the structure's dynamics using a low-modal-dimensional model, and so on

Note: DOF = degree of freedom.



**FIGURE 7** The experimental setup to perform video measurement of a three-story building structure excited by horizontal impact. (Accelerometers measurement were used in input–output testing for reference comparisons)

illumination enhancement. An impact hammer was used to excite the structure horizontally at different places for different testing scenarios, inducing only in-plane vibration. Figure 7 shows the experimental setup.

For validation, each “virtual” test (simulations using the established modal model) was compared with an “actual” test in the same setting. The following comparison method was used. For each experiment, the force applied on the structure in the actual test was measured and used in the simulations as the virtual force. This force was then projected to the modal model, and the induced simulated structural responses are obtained and embedded into the simulated video. These simulated structural responses are compared with the responses estimated from the video measurements of the “actual” test.

## 3.2 | Implementation procedures and results

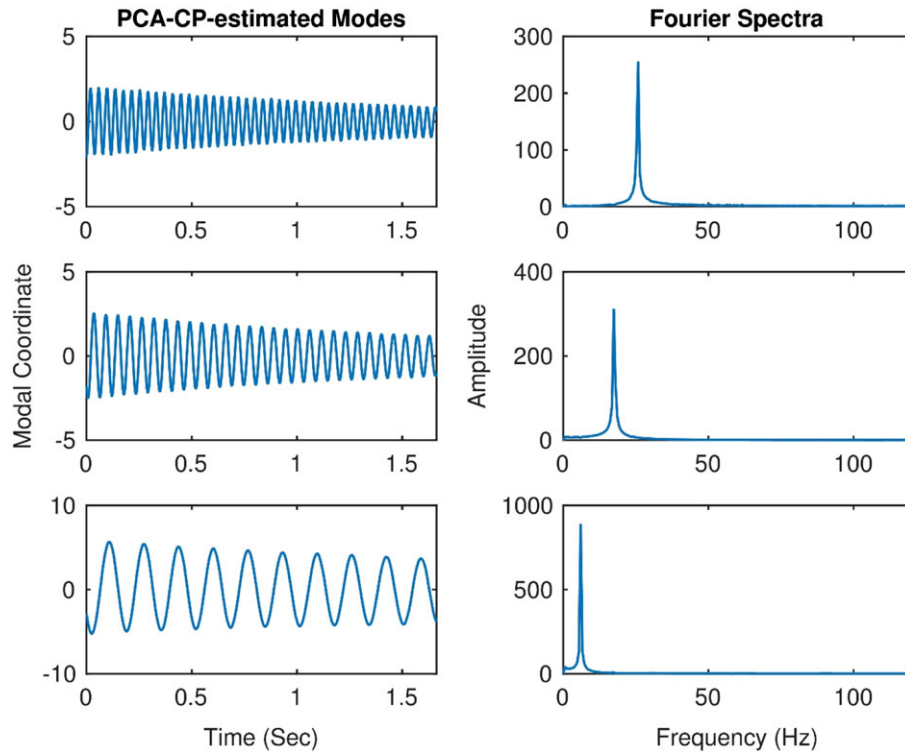
### 3.2.1 | Establishing the full-field modal model

For the first step, a hammer test was conducted to establish the full-field modal model. From the video measurements (the pixels were *spatially* downsampled to  $384 \times 216$  [with reduced spatial resolution but still very high spatial dimension for full-field measurement] to reduce the data points for the purpose of more efficient computation and better illustration/visualization of the results in the figures/videos in this experimental study\*), the displacement motion was extracted and analyzed by the recently developed video-based full-field output-only modal identification method<sup>[26]</sup> to identify the modal parameters. Figure 8 shows the extracted modal coordinates from which modal frequencies and damping ratios were then estimated by HT-based SDOF analysis method. Figure 9 shows the identified full-field mode shapes. The modal mass of each mode was estimated using the dominant mass of each floor and the mode shape component at the pixel coordinate corresponding to the floor location (Figure 10), according to Equation (24). Thus, the modal model was established.

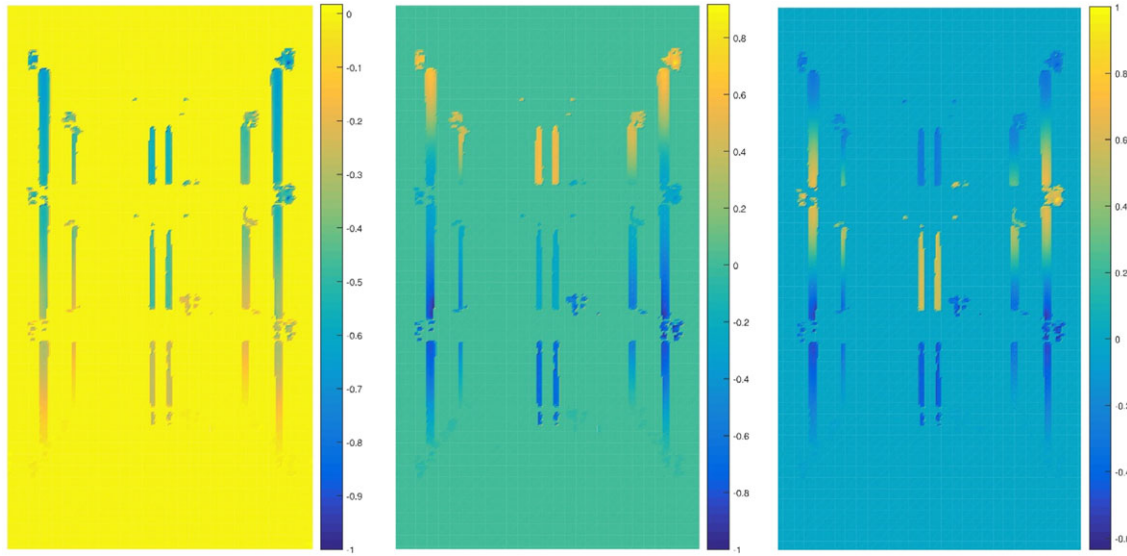
To validate the established modal model, a discrete-type FEM with three dominant masses was established, and additionally, an input–output hammer testing (producing the frequency response function) using accelerometers (contact type) measurements were conducted. Their estimated modal parameters are listed in Table 3, and the discrete-type mode shapes are shown in Figure 11. For calculating the modal assurance criterion values, the mode shapes from the FEM were used as reference. Also, the mode shape components of video-domain full-field modal model corresponding to the three locations of the three floors were correlated with the mode shapes of the FEM. It is seen that the modal parameters estimated from the videos have excellent correlations with others.

\*It is found in the experiments that the computation increases gracefully with the increase of the pixel dimension. Note that pixel downsampling is not a required step by the proposed method; more efficient optical flow methods<sup>[44]</sup> to extract full-field displacements from the videos may be explored for computational efficiency but is out of the scope of this study.





**FIGURE 8** The modal coordinates and their Fourier spectra extracted by the principal component analysis and complexity pursuit (PCA-CP) algorithm applied on the full-field spatiotemporal pixel phases of the video measurements

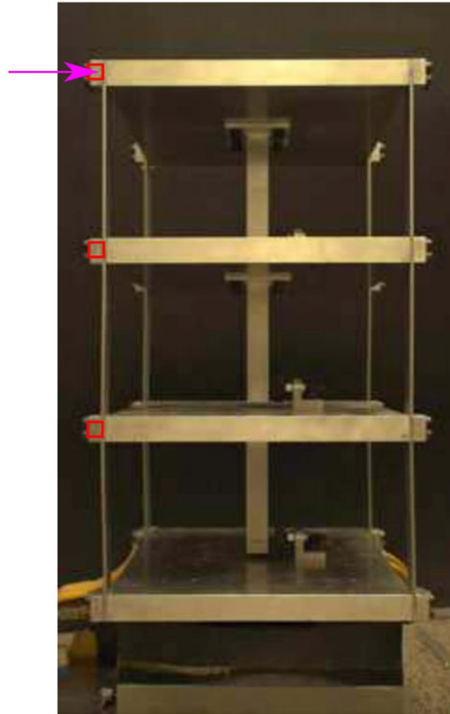


**FIGURE 9** The full-field mode shapes of the three-story structure in the pixel coordinates ( $x$  and  $y$  axis) identified by the principal component analysis and complexity pursuit algorithm applied on the full-field spatiotemporal pixel phases of the video measurements. From left to right: Mode 1, Mode 2, and Mode 3. Note that the rigid floors are not present. Also note that the color bar scales are different for each plot

### 3.2.2 | Video-domain simulation and visualization of full-field dynamic response

Then different impact force excitations were applied on the structure for different simulation scenarios: Test 1 with impact force on the first floor and Test 2 on the top floor, respectively. The responses (displacements) induced by the





**FIGURE 10** The pixel coordinates of the video frames of the three-story building structure. Red squares are the three pixel coordinates used for estimating the modal masses (each red square contains only one pixel; the enlarged square is only for enhanced visualization in this plot). The pink arrow points to the pixel coordinate where the (virtual) force is applied (this plot shows the top floor impact case) during the testing in the video-domain simulation

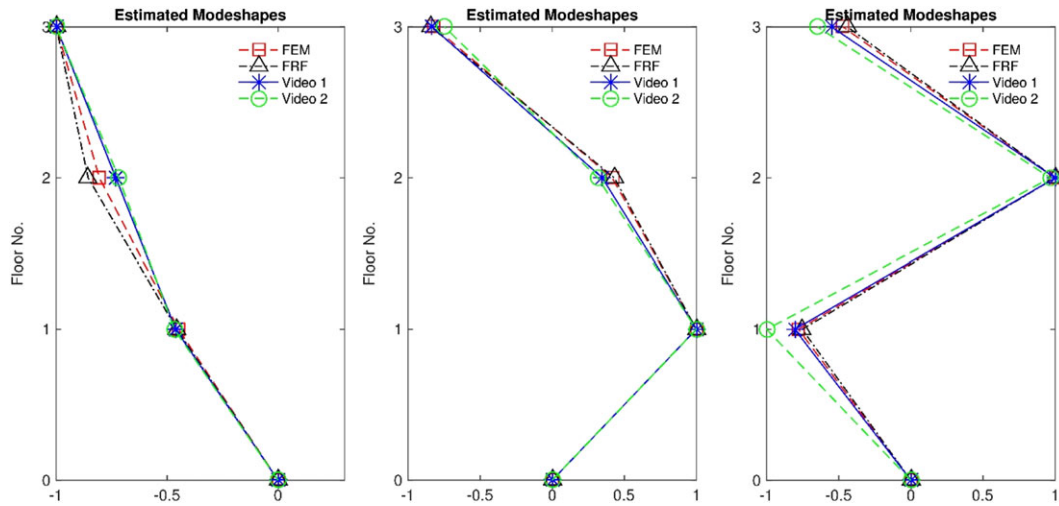
**TABLE 3** The modal parameters estimated from different approaches

Mode	Modal frequency (Hz)			Damping ratio (%)			MAC (discrete mode shape)		
	1	2	3	1	2	3	1	2	3
FEM	6.35	17.70	25.49	—	—	—	—	—	—
FRF	6.10	17.20	25.30	0.68	0.65	0.36	0.9992	0.9998	0.9995
Video S1	6.00	17.40	25.80	0.64	0.41	0.30	0.9976	0.9974	0.9973
Video S2	6.00	17.40	25.80	0.63	0.53	0.30	0.9966	0.9941	0.9773

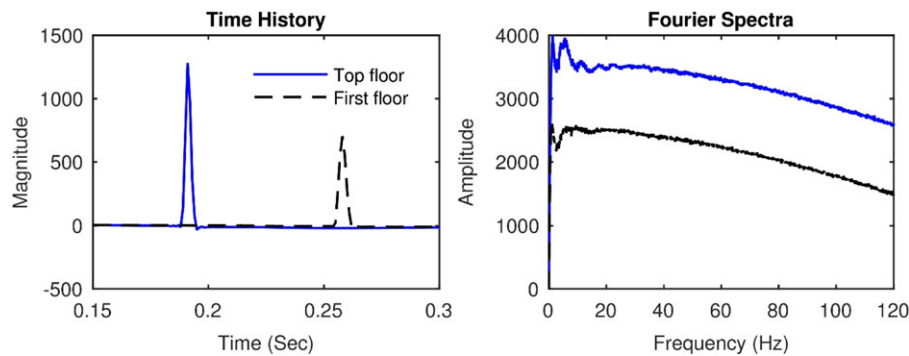
Note: FEM = finite element model; FRF = frequency response function; MAC = modal assurance criterion.

actual measured impact force (Figure 12) were simulated using the established modal model (solving Equation (21) by the Newmark-beta algorithm) and modal superposition synthesis Equation (25). To validate these simulated full-field responses (displacements), they were compared with those estimated from the video measurements of the “actual” test using the video-based output-only modal identification technique.<sup>[26]</sup> For demonstration, the responses of some pixels are shown in Figure 13 and Figure 14 for Test 1 and Test 2, respectively. Since time synchronization between the force measurement system and video cameras was not possible, the two types of responses were temporally aligned manually for comparisons. Also to focus on the comparisons of the dynamics of the simulated and the experimentally estimated structural response, the responses were normalized by their maxima, respectively, excluding the errors of the calibration factors of these two measurement systems. Note that the normalized responses have no physical units.

It is seen that the simulated responses correlate quite well with the motion estimated from the video measurements of the “actual” tests in both time domain and frequency domain. Time response assurance criterion (TRAC), similar to the frequency response assurance criterion,<sup>[45]</sup> was used to quantify such correlation. The TRAC values of the responses of Test 1 (for the three pixels shown in Figure 13) are 0.8811, 0.8005, and 0.8365, and those of the response of Test 2 (for



**FIGURE 11** The estimated mode shapes using three different approaches: red square— finite element model (FEM); black triangle— Input-output testing frequency response function (FRF); blue star (Test 1) and green circle (Test 2)—video processing technique. The x axis is the normalized magnitude. Note that the mode shapes estimated from the videos were full-field (Figure 9), but their values at three locations corresponding to the three floors are plotted here to compare with the other two approaches



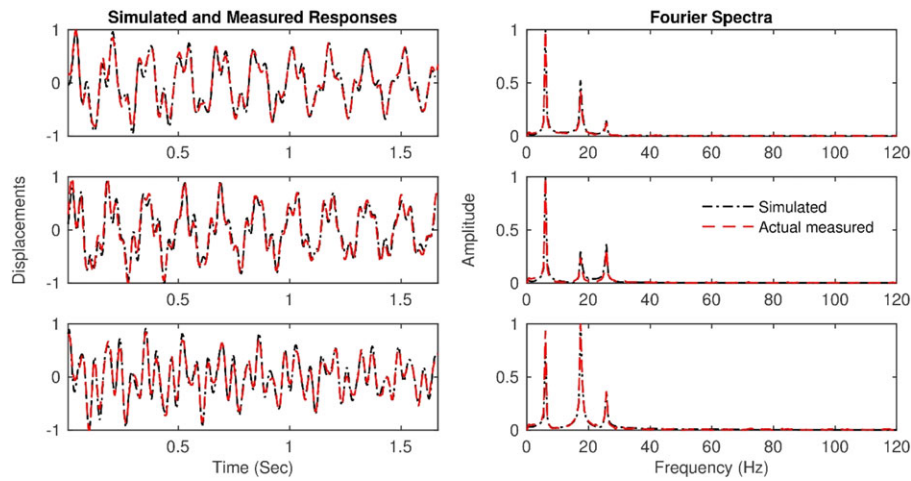
**FIGURE 12** The time histories and their Fourier spectra of the hammer impact forces applied on the three-story building structure when actually conducting the tests for comparisons with the video-based motion simulation and synthesis. (blue solid line: impact applied on top floor [Test 2]; black dash line: impact applied on the first floor [Test 1])

the three pixels shown in Figure 14) are 0.9580, 0.9530, and 0.9467. The high TRAC values indicate the effectiveness of the established modal model in characterizing the dominant structural dynamics and predicting dynamic responses.

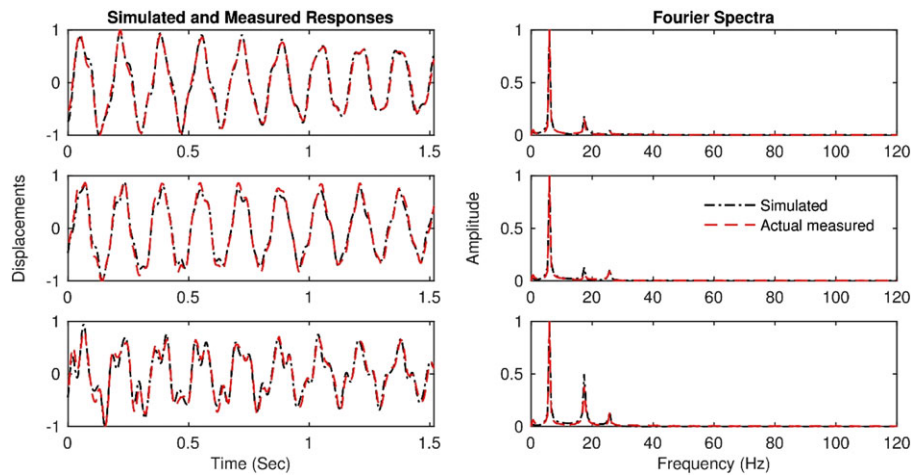
Finally, these simulated full-field responses were embedded into the original video by replacing the existing motion in the original video with these simulated motion. Video S3 and Video S4 are the simulated videos for the two force excitation cases, while Videos S1 and S2 are the video measurements of these two “actual tests.” Videos S1–S4 are provided in the Supporting Information section online. The good correlation of the motions evident in Figure 13 and Figure 14 implies that the embedded simulated motion of the generated video (Videos S3 and S4) were consistent with the motion of the real-recorded videos of the actual tests (Videos S1 and S2, respectively). This is visualized by comparing Video S3 with Video S1 and Video S4 with Video S2, respectively, where the simulated videos have very high resemblances to the actual videos. Note that between Videos S1 and S3 and between Videos S2 and S4 were not time synchronized, respectively. Therefore, it is validated that the proposed method enables high-fidelity simulations and realistic visualization of structural dynamic response.

### 3.3 | Discussions of the sources of errors and limitations for practical implementations

Although the results presented in Section 3.2 for the simulated response using the proposed method and the response estimated from the actual experimental tests are reasonably accurate in terms of the correlations quantified by TRAC,



**FIGURE 13** The synthesized structural responses (impact on first floor for Test 1) and their Fourier spectra using the full-field modal model of the three-story building structure, compared with those responses measured and estimated from the video measurements of the actually-conducted test. Showing only the (normalized) displacements of the three pixels (from top to bottom) circled on Figure 10. (black point dash line: synthesized; red dash line: actual) See the simulated full-field structural responses in Video S3 compared with the actual structural responses in Video S1



**FIGURE 14** The synthesized structural responses (impact on top floor for Test 2) and their Fourier spectra using the full-field modal model of the three-story building structure, compared with those responses measured and estimated from the video measurements of the actually-conducted test. Showing only the (normalized) displacements of the three pixels (from top to bottom) circled on Figure 10. (black point dash line: synthesized; red dash line: actual.). See the simulated full-field structural responses in Video S4 compared with the actual structural responses in Video S2

for the experiment in this study and for practical implementations, the errors of the simulated motion embedded to the new videos could come from several sources, which are summarized in Table 1 and detailed in the following with discussions on the experimental results.

### 3.3.1 | Main error sources

The first error source comes from the imperfect estimation of the full-field modal model from the video measurements. It is seen from Table 1 that the modal parameters among different estimation approaches are slightly different. First, the modal parameters estimated from the video, from which the full-field modal model was established, are different from those from input-output testing. This is because the modal parameters estimated from the video use output-only

response and are usually less accurate than those estimated using both input and output information. Also, both estimated from the videos (output-only response) but at different impact excitation scenarios, the modal parameters are also slightly different, especially the damping ratios of the second mode (Table 1) and the mode shapes of the third mode (Table 1 and Figure 11). This may be due to (weakly) nonlinear behaviors of the structure where varying excitation such as different amplitudes will induce change of dynamic properties of the structure. Therefore, it is seen that using the modal model estimated from the original video to simulate new dynamic response induced by different excitation causes difference between the simulated response and the actual response (see Figure 13 and Figure 14).

Another error source associated with the full-field modal model may be due to the discrepancy that, the new dynamic response was simulated using the established modal model with proportional damping (using the damping ratios estimated from the original video shown in Table 1), while the damping of the structure is usually nonproportional. This could cause the slight difference between the simulated and the actual decaying behaviors of the structural response as seen in Figure 13 and Figure 14. In addition, it is a common challenge in operational or output-only modal analysis that higher modes are less sufficiently excited, and their estimation is less accurate (e.g., the damping ratio of the second mode and the third mode shape estimated from the video, see Table 1). The higher modes in the modal model are thus less accurate. This is seen by comparing the simulated responses of the two tests, shown in Figure 13 and Figure 14. The response of Test 1 contains more components of Mode 2 and Mode 3 (in which the established modal model is less accurate), and the simulated response has less correlation (lower TRAC values) with the actual response (Figure 13). On contrary, the response of Test 2 has dominant Mode 1 component (in which the established modal model is more accurate) and has higher correlation (higher TRAC values) with the actual response (Figure 14). Another reason may be due to that the higher amplitude of the impact force in Test 2 than Test 1 induces higher magnitudes of the structural responses and thus has higher signal-to-ratio of the measurements.

### 3.3.2 | Issues and limitations for practical implementations on real structures

Finally, the deviation between the actual motion direction and the projected motion in the camera (lens) perspective (video coordinates) could also induce error. In practice, this error could be reduced to minimal by more accurate camera setup and calibration strategy. Also, real structures usually exhibit both in-plane and out-of-plane (three-dimensional) vibration; it is therefore necessary to extend the proposed method to also estimate out-of-plane motion (three-dimensional, e.g., torsional modes). This can be accomplished using the multiscale (including multioriented) complex *steerable* pyramid filters of the phase-based optical flow method in a three-dimensional formation to allow for estimation of in-plane and out-of-plane motion in any orientation with *stereo* camera measurement system, while this study focuses on proof-of-principle development and validation of the proposed algorithm with the basic one-dimensional in-plane vibration cases.

For possible solutions in the future to improve the method, a modal model additionally incorporating nonlinear modeling terms could help address this issue; challenge exists for identifying and modeling the nonlinearity, however. Also, as discussed in the Estimation of the Modal Mass Section, the estimation of modal masses is approximately applicable for structures with rigid motion of dominant masses; otherwise, a full mass matrix is needed but subjected to availability. In addition, more robust algorithms are needed for identifying the full-field modal model from the video, as video camera measurements of outdoor real-world structures have varying imaging environment, such as variant perspective, illuminations, and other environmental effects.

## 4 | CONCLUSIONS AND FUTURE WORK

This study proposes a novel video-based method for high-fidelity simulation and realistic visualization of full-field structural dynamic response based on a combination of the full-field modal model and the video motion manipulation techniques. The full-field modal model is established by estimating the full-field modal parameters from line of sight video measurements of the structural vibration using the phase-based video motion estimation and BSS-based full-field output-only modal identification algorithm. The modal masses are approximated by a simple model reduction strategy. The modal responses simulated from the modal model subjected to the modally-projected arbitrary force are synthesized by the full-field mode shapes using modal superposition to obtain the full-field structural response. The simulated full-field structural response is finally embedded to the original video of the structure to render full-field, high-fidelity, realistic visualization, and animation of the simulated motion in a realistic video of the structure. This new method is

validated on a proof-of-principle basis by simulation study on a numerical model and experimental study on a laboratory-scale model of a building structure.

The proposed method is potentially a useful alternative to the traditional finite element computer model simulation method in alleviating the exhaustive procedures and computation costs associated with high-fidelity modeling and simulating structural dynamic behaviors. The first reason is because the method bypasses the need for explicit modeling of the materials properties, structural geometry, and boundary conditions. The second reason is that the method directly exploits the low-modal-dimensional nature of structural dynamics. An additional potentially attractive feature of the method is that it enables high-resolution visualization and animation of the simulated full-field structural dynamic response in a realistic video of the structure that was not previously possible. Nevertheless, there are a few issues that need to be addressed before the proposed method could possibly gain wider acceptance. Future work needs to be conducted on building more accurate dynamic model by possibly conducting more refined experimental modal analysis and incorporating nonlinear terms and extending the dynamic model and motion processing and synthesis to three-dimensional using stereo video camera measurements and processing for real-world structures.

## ACKNOWLEDGEMENTS

We would like to acknowledge the support of the Los Alamos National Laboratory Lab Directed Research and Development (LDRD) program. This program has supported this work in the form of a Director's funded postdoctoral fellowship (Grant 20150708PRD2) for Yongchao Yang. Part of the work was conducted during the 2015 Los Alamos Dynamics Summer School.

## ORCID

Yongchao Yang  <http://orcid.org/0000-0003-1776-3306>

## REFERENCES

- [1] J. N. Reddy, *An Introduction to the Finite Element Method*, McGraw-Hill, New York **2005**.
- [2] M. Friswell, J. Mottershead, *Finite element model updating in structural dynamics*, Springer, Netherlands **1995**.
- [3] C. R. Farrar, K. Worden, *Structural health monitoring: a machine learning perspective*, John Wiley & Sons, New York **2012**.
- [4] D. Ewins, *Modal testing: theory, practice and application*, Research Studies Press LTD, Baldock, Hertfordshire, England **2000**.
- [5] Y. Yang, S. Nagarajaiah, *Mech. Syst. Sig. Proc.* **2016**, 74, 165.
- [6] R. Brincker, C. Ventura, *Introduction to operational modal analysis*, John Wiley & Sons, New York, NY **2015**.
- [7] A. Stanbridge, D. Ewins, *Mech. Syst. Signal Proc.* **1999**, 13(2), 255.
- [8] M. Sutton, J. Ortu, H. Schreier, *Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications*, Springer Science & Business Media, New York, NY **2009**.
- [9] B. Horn, B. Schunck, *Artifi. Intell.* **1981**, 17, 185.
- [10] A. Wahbeh, J. Caffrey, S. Masri, *Smart Mater. Struct.* **2003**, 12(5), 785.
- [11] C. Chang, Y. Ji, *J. Eng. Mech.* **2007**, 133(6), 656.
- [12] Y. Fukuda, M. Feng, M. Shinozuka, *Struct. Control Health Monit.* **2010**, 17(8), 918.
- [13] H. Yoon, H. Elanwar, H. Choi, M. Golparvar-Fard, B. Spencer Jr., *Struct. Control Health Monit.* **2016**, 23(2), 1405.
- [14] Z. Chen, H. Li, Y. Bao, N. Li, Y. Jin, *Struct. Control Health Monit.* **2016**, 23(3), 517.
- [15] Y. Chen, M. Abdelbarr, M. Jahanshahi, S. Masri, *Struct. Control Health Monit.* **2017**, 24(11), e2000.
- [16] T. Siebert, R. Wood, K. Splithof, *J. Phys. Conf. Ser.* **181**(1, P. 012064), 2009.
- [17] M. N. Helfrick, C. Niezrecki, P. Avitabile, T. Schmidt, *Mech. Syst. Signal Process.* **2011**, 25(3), 917.
- [18] W. Wang, J. Mottershead, J. Ihle, T. Siebert, H. Schubach, *J. Sound Vib.* **2011**, 330(8), 1599.
- [19] W. Wang, J. Mottershead, T. Siebert, A. Pipino, *Mech. Syst. Signal Process.* **2012**, 28, 333.
- [20] J. Baqersad, C. Niezrecki, P. Avitabile, *J. Sound Vib.* **2015**, 352, 16.
- [21] J. Javh, J. Slavič, M. Boltežar, *Mech. Syst. Sig. Process.* **2017**, 88, 89.
- [22] D. Fleet, D. Allan, *Int. J. Comput. Vision* **1990**, 5(1), 77.



- [23] N. Wadhwa, M. Rubinstein, F. Durand, N. Wadhwa, *ACM Trans. Graph.* **2013**, 32(4), 1.
- [24] T. Gautama, M. Van Hulle, *IEEE Trans. Neural Netw* **2002**, 13(5), 1127.
- [25] J. Chen, N. Wadhwa, Y. Cha, F. Durand, W. Freeman, O. Buyukozturk, *J. Sound Vib.* **2015**, 345, 58.
- [26] Y. Yang, C. Dorn, T. Mancini, Z. Talken, G. Kenyon, C. Farrar, D. Mascarenas, *Mech. Syst. Sig. Process.* **2017**, 85, 567.
- [27] O. Arikan, D. Forsyth, *ACM Trans. Graph. (TOG)* **2002**, 21(3), 483.
- [28] M. Sun, A. Jepson, E. Fiume, in *Proceedings of Ninth IEEE International Conf. on Computer Vision*, **2003**.
- [29] A. Davis, J. Chen, F. Durand, *ACM Trans. Graphics (TOG)* **2015**, 34(6), 239.
- [30] J. Portilla, E. Simoncelli, *Int. J. Comput. Vis.* **2000**, 40(1), 49.
- [31] B. Feeny, R. Kappagant, *J. Sound Vib.* **1998**, 211, 607.
- [32] J. Antoni, *Mech. Syst. Signal Process.* **2005**, 19, 1166.
- [33] G. Kerschen, F. Poncelet, J.-C. Golinval, *Mech. Syst. Signal Process.* **2007**, 21, 1561.
- [34] S. McNeill, D. Zimmerman, *Mech. Syst. Signal Process.* **2008**, 22(7), 1526.
- [35] A. Sadhu, B. Hazra, S. Narasimhan, *Mech. Syst. Signal Process.* **2013**, 41(1), 396.
- [36] Y. Yang, S. Nagarajaiah, *Earthq. Eng. Struct. Dyn.* **2013**, 42(13), 1885.
- [37] S. Ghahari, F. Abazarsa, M. Ghannad, E. Taciroglu, *Earthq. Eng. Struct. Dyn.* **2013**, 42(8), 1221.
- [38] J. Antoni, S. Chauhan, *J. Sound Vib.* **2013**, 332(4), 1079.
- [39] P. Brewick, A. Smyth, *J. Sound Vib.* **2014**, 333(26), 7333.
- [40] Y. Guo, A. Kareem, *J. Sound Vib.* **2016**.
- [41] S. Nagarajaiah, Y. Yang, *Struct. Control Health Monit.* **2017**, 24(1), e1851.
- [42] A. Agnani, *Int. J. Anal. Exp. Modal Anal.* **1992**, 7(4), 285.
- [43] "PEER Ground Motion Database," [Online]. Available: <http://ngawest2.berkeley.edu/>.
- [44] R. Allemang, *Sound Vib.* **2003**, 37(8), 14.
- [45] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, R. Szeliski, *Int. J. Comput. Vis.* **2011**, 92(1), 1.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Yang Y, Dorn C, Mancini T, et al. Spatiotemporal video-domain high-fidelity simulation and realistic visualization of full-field dynamic responses of structures by a combination of high-spatial-resolution modal model and video motion manipulations. *Struct Control Health Monit.* 2018;25:e2193. <https://doi.org/10.1002/stc.2193>